

NEW "ELECTRONIC" DATA COLLECTION METHODS in the NATIONAL AUTOMOTIVE SAMPLING SYSTEM CRASHWORTHINESS DATA SYSTEM

Seymour D. Stern
Marvin M. Stephens, Jr.
Kenneth W. Rutland

National Highway Traffic Safety Administration
United States
Paper Number 98-S6-P-18

ABSTRACT

The National Automotive Sampling System (NASS) Crashworthiness Data System (CDS) was established in 1979 to help analyze motor vehicle crashes and injuries. It collects data with research teams at 24 sites that study about 5,000 crashes annually involving passenger cars, trucks, and vans. Historically, trained researchers have collected NASS CDS data on paper data encoding forms.

Recently, the NASS program took a giant step forward into the world of electronic crash data collection. Researchers started collecting data using electronic digital methodologies in January of 1997. In this "paperless" system data go directly into the database in the field, rather than on paper forms.

All case data are transferred electronically to quality control centers, central data depositories, and NHTSA. An interface is being designed for users to access data and digital images. This paper will describe the process of converting from a paper to an electronic crash data collection system. This new system gives NHTSA the capability to provide quick electronic access to cases. Eventually the data will be accessible through the Internet.

INTRODUCTION

In 1991, NHTSA first began the process of investigating the possibility of scanning NASS CDS cases to provide easy access for clinical analysis. This was followed with an investigation into the current state-of-the-art in still video and digital cameras. In 1995, NHTSA reassessed the future needs and began design work on a total electronic system for acquiring crash data, not just a conversion of conventional data collection efforts. Objectives were set for an electronic data file to be made available to the public in late 1998 containing data, images, and photographs for crashes occurring during calendar year 1997.

Increasing the accessibility and timeliness of Agency data was part of Goal 4 of the Strategic Execution Plan for the agency. An objective of this goal was to "develop software for user-friendly, direct access to NHTSA's major data files through the Internet and other electronic media." This was set as a milestone for the year 1999. The new NASS electronic data collection system is the vehicle that will allow the agency to achieve that milestone and continue to provide responsive and timely service to data users throughout the United States and the world.

HARDWARE / SOFTWARE

Field Computers

In June 1996, the NASS development team purchased the hardware that would become the electronic "tools" necessary to implement the paperless data collection system. Fujitsu Stylistic 1000 pen-based computers were purchased for all field researchers, for the NASS training instructors, for NASS headquarters and for the development team. A depot of spare units is maintained to replace units out of service.



Figure 1. This is the Fujitsu Stylistic 1000 pen-based computer that is used for data collection.

Digital Cameras

To replace the old system that recorded information on 35 mm slides, the development team began in early 1992 to look at alternative imaging techniques. Initially it was thought that still video imaging would be the best method. However, the advent of the digital camera around 1995 provided a convenient format to capture and record electronically the data that had traditionally been gathered through 35 mm slides. The number of digital cameras available at that time was limited. After reviewing the available models, the development team chose the Fujix DS-220 based on cost, color resolution, its wide angle/close up features, its LCD screen preview, and its available PC card storage media. The digital cameras were purchased at the same time as the pen-based computers, and each researcher and other team members who had a pen-based computer also had a camera. A

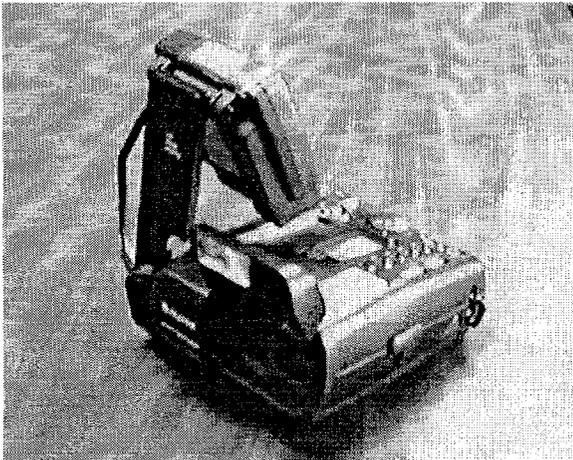


Figure 2. The Fujix DS-220 digital camera with LCD monitor attached.

depot of spare cameras is also maintained.

Because the technology continues to improve in the computer and camera marketplace, the NASS team is continuously reviewing the available products so that when the system hardware needs replacing, appropriate replacements can be specified. Laptops and pen-based computers are being considered, and the latest digital camera technologies that permit increasingly more sophisticated digital imaging and storing possibilities are being reviewed.

Data Storage

Gathering of the data electronically was only half of the

equation. Plans had to be made for storage and retrieval of vastly larger amounts of data, larger than the previous data collection efforts. The average annual number of cases collected in the NASS did not change from the previous system to the new 1997 start of the electronic data collection methodology (5,000). However, the data to be stored increased, specifically the number of variables and the number and size of the digital images recorded. Previous electronic "flat" file storage per case was 3 kilobytes or approximately 15 megabytes per year. With the new electronic data collection and storage, the text and digital images total approximately 30 megabytes per case or 125 gigabytes for the year. Consequently, a relational database was required to ensure that the data variables and all associated images could be stored and easily retrieved. Therefore, an Oracle database was developed and several sub programs were developed for getting information from the pen-based computers into the database. For example, programs were developed to enter the Police Accident Report (PAR) data into the main program, to enter the field data collected by the researcher, to import and organize the digital images, to create scaled scene diagrams, to import scanned images, to provide for injury coding, and to permit monitoring and data management.

Work continues developing mechanisms to provide continuous data access to analysts and researchers who have in the past used the NASS Crashworthiness Data System (CDS) and General Estimates System (GES) data, and new groups of users. NHTSA has implemented a DARS (Data Archival and Retrieval System) group to test and evaluate methods of using the data from the relational database. Results of this test and the programs being designed by the development team are expected to bring the data to the public in the summer of 1998. The data collection will continue to be upgraded and improved in future NASS electronic data collection years.

PSU TEAM FUNCTIONS

Sampling

The basis for the NASS is its sampling algorithm by which each Primary Sampling Unit (PSU), or research team, lists each PAR and selects the appropriate ones for further investigation. A team member goes to the jurisdiction and lists information for each valid NASS CDS PAR filed since their last visit. This was previously done using paper and pencil. The researcher then returned to their office and entered this information into a desktop computer where it was then uploaded to another computer.

Now the information is entered at the police jurisdiction into a pen-based computer. A check for errors and duplicate entries is made at the time of the original listing (instead of later at the office). When the listing is complete, the researcher returns to the office, uploads the information into the local computer system and sends it through a Frame Relay connection to a computer at our software support contractor where the sample is selected and down loaded.

The police report for the selected crash is obtained and scanned. The scanned report is used for quality control and kept separate from the case information.

Scene Documentation

The crash scene is required to be documented, photographed, and have a scene diagram drawn for each crash researched. Photographs and scene diagrams have always used conventional 35mm slides and paper and pencil diagrams. We now use a Fujix DS-220 digital camera for all photography. This camera permits previewing and reviewing all photos. The digital images are all stored on PC cards and uploaded into the pen-based computer back at the office where they are categorized and labeled. Rough sketches of the scene are done in the field. These documents are then used to produce Visio® diagrams of the scene at the office. The field sketches are scanned and included as case documentation.

Coding Of Data

As with sampling, previous procedures involved coding data on paper forms in the field and returning to the office to enter the information into the computer. The pen system is now used for direct field entry of the information. It especially facilitates the sketching of vehicle damage on an outline of the vehicle on the pen-based computer. The field researcher is led through the process of entering data and can never run short of forms. This information is then uploaded back at the office into the local system where edit checks on the data can be run anytime. Photos of the car are taken with the digital camera. The camera permits the researcher to photograph out-of-the-reach locations and preview the image. Its instant feedback of information produces increased opportunities for quality control and assurance of adequate photographic coverage.

Interviews

Interviews are still recorded on paper forms. When interviews are conducted in person, the paper can be less intimidating than a computer to the person being interviewed. Paper also simplifies the rapid documentation of information even when the interview is conducted over the phone. The completed form is then scanned and included with the case materials.

Medicals

Official medical records are obtained and shipped to the quality control zone centers. This information is recorded and coded by experts. Afterwards, the medicals are scanned and the originals destroyed. The originals are used for this coding process as the quality of the handwriting has always been an issue.

ZONE CENTER FUNCTIONS

Coding Of Injury Data

The "zone centers" are the quality control center for a set of PSUs. Their expertise permits them to code the injury data for all crash involved persons. They review the hard copy of the medical report received from the PSU, placing injury codes and descriptions directly on mannequins of the human body at the place of injury displayed on the computer screen. This permits not only an automated review of the data coded, but also a clinical review of the injuries and their locations on the body. After the injuries have been coded and checked, the hard copies of the medical reports are destroyed, having already been scanned at the relevant PSU.

Case Quality Control

The zone center reviews each case and its components at different times during the case review process. Corrections to the case are made directly on-line and reasons for the change are documented at the same time. This permits summary reports and tracking of the progress and problems for individual researchers and teams.

Productivity and Timeliness

The zone centers also use the system to track the timeliness and productivity of interviews, scene and vehicle inspections, and case throughput. The date of the interview, inspection, etc., is automatically recorded in the system, where previously it had to be manually entered. This permits a rapid and up-to-date review of the progress

of all field personnel.

Case updates are made each night from the PSU to the zone center, to the central files in Cambridge Massachusetts at the software contractor, and to NHTSA offices in Washington, D.C. This permits the review and monitoring of any case at anytime.

Expertise

The zone centers are considered our experts in field crash data collection. If a researcher has a question about data coding, collection, or crash reconstruction the zone center now has the capability of looking directly at the same case. This permits an easier and more helpful way of aiding experienced researchers who encounter a confusing situation, a novice researcher who has not had much field experience, and any researcher who needs help with new procedures, protocol, or variables.

NHTSA FUNCTIONS

Quality Control

NHTSA regularly examines the file looking for ways to improve the data. Previously when a case was found where data were in question and the case had already been finished and shipped to Washington, D.C. (for storage), the case had to be shipped back to the zone center for review and modification. Now the zone center or NHTSA can review the case anytime and verify it for correctness.

All cases are available to quality control staff at NHTSA anytime in the process. In the old system they were not available until months after data collection was completed and they were shipped to Washington, D.C.

Case Access

In the new "electronic" environment, NHTSA users will have several ways to use the system. They can access it from their desktop computers (when set-up for this connection) and query the system using either SAS, Oracle SQL queries, Crystal Reports, or the Data Archival and Retrieval System (DARS). The DARS system is still under development, but will ultimately permit access to cases and viewing of their contents.

OVERALL BENEFITS OF SYSTEM

The system has been designed to deal with several

separate issues, therefore benefits fall into several categories.

No Paper

This reduces the need for mailing, envelopes, handling, shipping, storage and retrieval of cases. This reduces cost through the elimination of printing, duplication, postage, and handling of cases for purposes of their storage and retrieval (for requests received). The introduction of direct field entry of the data eliminates double data entry and the inherent problems of incorrect entry of the field data from the forms.

No 35-mm Slide Photography

This eliminates film purchase, developing costs, slide sorting, and slide duplication when a copy of a case is requested. With the use of digital photography, a photo can be duplicated anytime. However, the quality of the image is not as good as that of a 35-mm slide transparency. This will improve as we acquire higher resolution cameras in the future.

Cases will be available for NHTSA researchers directly from their desktop computers

This will eliminate the process of having to visit the case storage facility to view cases. Eventually this type of access may be available to the public via the Internet.

General Vehicle Form, Case 939/ Vehicle #1

Vehicle | Official Records | Pre Crash Environment | DRIVER | AOPS | Rollover | Reconstruction | Delta V | LOG

Precrash Environmental Data

General		Conditions
Interchange or Junction	Interchange area related	Light
Trafficway Flow	Divided trafficway-median strip with c	Atmospheric
Travel Lanes	Two	
Roadway		Traffic Control
Alignment	Curve right	Device
Profile	Uphill grade (>2%)	Functioning
Surface Type	Brick or black	
Surface Condition	Wet	

OK Cancel

Figure 3. The pre-crash environmental data entry screen that researchers use for entering data.